Assignment 4

NAME – Vishal Kumar

Roll No – 2018IMT-106

Course – Machine Learning (ITIT 4103)

Deadline – 18th October 2021

Aim

1. Use a scatter plot to visualize the data, since it has only two properties to plot (profit and population).

2. Consider a simple linear model with two parameters and one input variable and mean square error cost function to implement the gradient descent algorithm to find the intercepts. Assume a suitable terminating condition.

3. Plot the model alongside the scatterplot to show the fit model.

4. Perform steps 1,2,3 in batch mode for varying values of alpha, learning rate and plot the results.

5. For each of the experiments performed above in steps 1,2,3,4 with varying learning rates visualize the cost function as a contour plot as well as plot the values of parameters to visualize the stepwise traversion of the parameters on this contour plot.

- Read data using panda
- Plot x and y –

Population vs Profit

- Mean square error cost

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)^2$$

- Linear regression model

$$h_\theta(x) = \theta^T x = \theta_0 + \theta_1 x_1$$
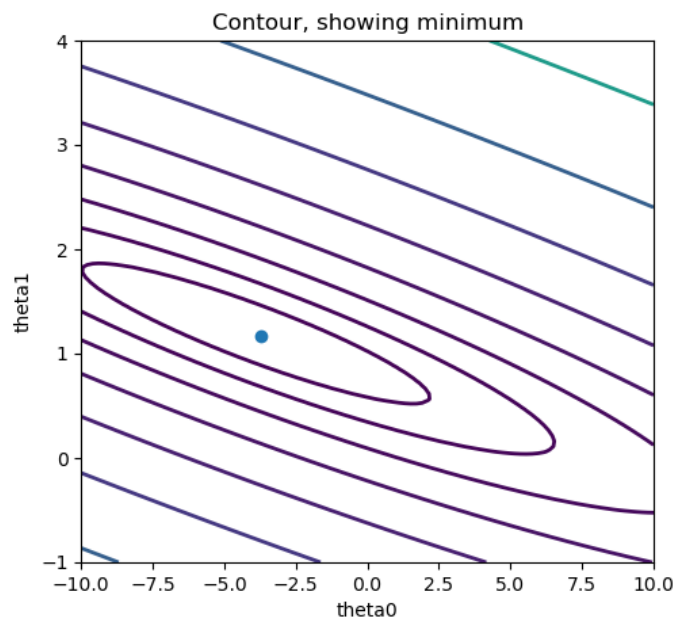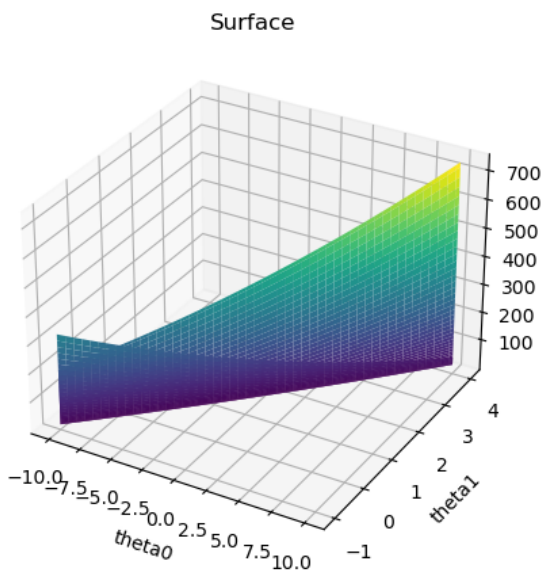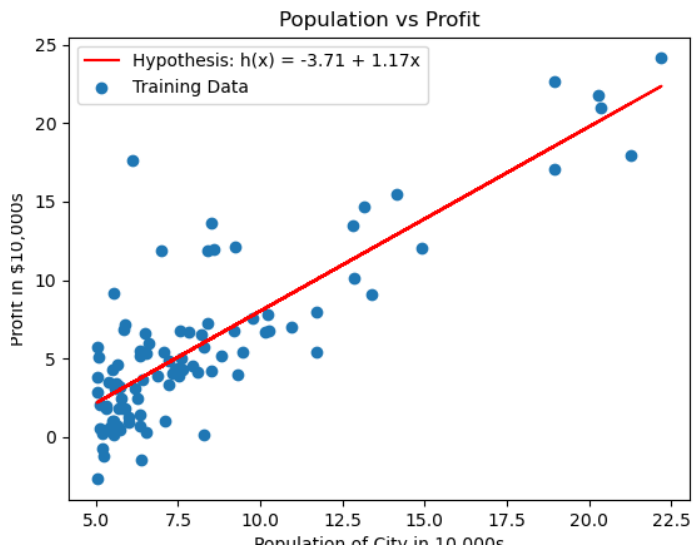
- We build gradient descent model using the equation

$$\theta_j = \theta_j - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right) x_j^{(i)} \quad \text{simultaneously update } \theta_j \text{ for all } j$$
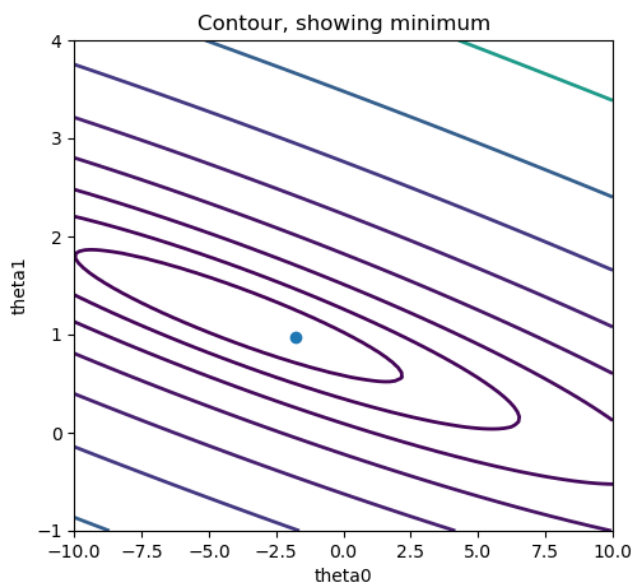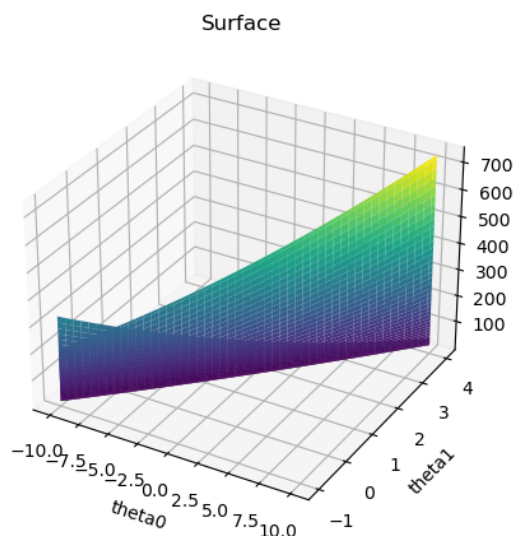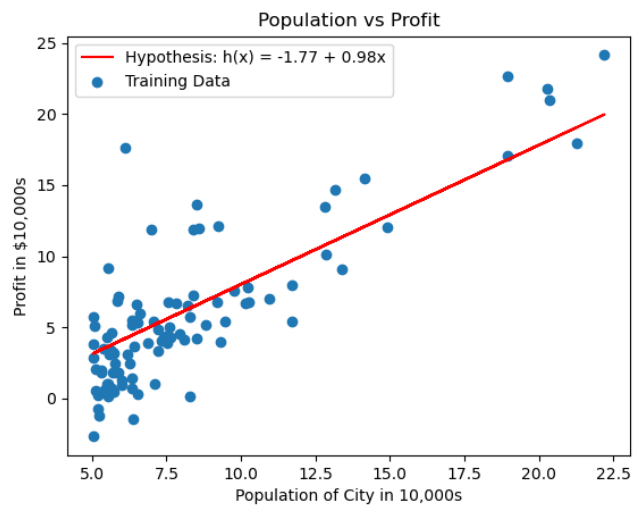
Do this for titao tita1 for n steps n=1500

- Contour built using matplotlib plt contour function
- For tita 0 from range -10,10 and tita 1 from range -1,4 cost is calculated and these 3 variables are passed to contour function

Result

Learning rate 0.01

Population vs Profit

Surface

Contour, showing minimum

Learning rate 0.001

Population vs Profit

Hypothesis: h(x) = -1.77 + 0.98x
Training Data

Surface

Contour, showing minimum

Learning rate 0.0001

Population vs Profit

Hypothesis: h(x) = -1.18 + 0.92x
Training Data

Surface

Contour, showing minimum